UCRL-CONF-204537

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Convergent Coarseness Regulation for Segmented Images

David W. Paglieroni

June 7, 2004

IGARSS04
Anchorage, AK, United States
September 20, 2004 through September 24, 2004

Convergent Coarseness Regulation for Segmented Images [1]

David W. Paglieroni, Senior Member IEEE

Lawrence Livermore National Laboratory
P.O. Box 808, L-290, Livermore, CA 94550
email: paglieroni1@llnl.gov vmail: 925-423-9295

Abstract

In segmentation of remotely sensed images, the number of pixel classes and their spectral representations are often unknown a priori. Even with prior knowledge, pixels with spectral components from multiple classes lead to classification errors and undesired small region artifacts. *Coarseness regulation* for segmented images is proposed as an efficient novel technique for handling these problems. Beginning with an over-segmented image, perceptually similar connected regions are iteratively merged using a method reminiscent of region growing, except the primitives are regions, not pixels. *Interactive* coarseness regulation is achieved by specifying the area α of the largest region eligible for merging. A region with area less than α is merged with the most spectrally similar connected region, unless the regions are perceived as spectrally dissimilar. In *convergent* coarseness regulation, which requires no user interaction, α is specified as the total number of pixels in the image, and the coarseness regulation output converges to a steady-state segmentation that remains unchanged as α is further increased. By applying convergent coarseness regulation to AVIRIS, IKONOS and DigitalGlobe images, and quantitatively comparing computer-generated segmentations to segmentations generated manually by a human analyst, it was found that the quality of the input segmentations was consistently and dramatically improved.

1. Introduction

Segmentation of images into regions containing pixels that logically belong together is an important early step in image analysis and interpretation. This paper uses a three-stage approach that involves pre-processing followed by conventional segmentation and post-processing. The goal of pre-processing is to produce an image suitable for both manual and automated interpretation. Typical pre-processing techniques include range clipping / quantization (for brightness-contrast adjustment), despeckling ([1]), and luminance-to-brightness mapping (to compensate for nonlinearities between sensed luminance and perceived brightness).

Most segmentation research has been focused on development of different families of image segmentation algorithms. These can be broadly categorized as pixel classifiers (e.g., [2]-[3]), region growers (e.g., [4]), edge-based methods (e.g., [5]) and evolving contour methods (e.g., [6]-[7]). These algorithms mostly require one or more input parameters for which appropriate values are unknown in advance. This impasse to fully automated processing is addressed in this paper by over-segmenting the images (which is easy to do), and then subjecting the results to a novel process referred to herein as *coarseness regulation* (discussed in Section 2). A robust method for image segmentation quality assessment is reviewed in Section 3 and then used in Section 4 to compare image segmentation quality before and after coarseness regulation.

[1] This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory, under Contract No. W-7405-Eng-48.

2. Coarseness Regulation

Coarseness regulation seeks to eliminate regions with less than α pixels by merging each such region with the most spectrally similar connected region. However, regions are not merged if the spectral difference between them exceeds Δ_s. Merged region i can be characterized by the set $r_i \stackrel{\Delta}{=}$ {ID's of unmerged regions contained in merged region i}. Initially, there are n unmerged regions, $r_i = \{i\}$ for $i = 0 \ldots n-1$, and the ID m_i of the merged region that contains unmerged region i is $m_i = i$. Let r_j be the most spectrally similar region connected (typically 4-connected) to region r_i. The spectral difference between r_i and r_j is taken to be $\| \mathbf{s}(r_i) - \mathbf{s}(r_j) \|$, where $\mathbf{s}(r_i)$ is the "spectral signature" of region r_i (i.e., the mean of pixel spectra for pixels in r_i). The sets $c(r_i)$ of ID's of merged regions connected to r_i are initialized by connectivity analysis on the initial set of unmerged regions. Coarseness regulation can be summarized as follows:

for $i = 0 \ldots n-1$

1. skip r_i if previously absorbed or too large:

   if $m_i \neq i$ or $area(r_i) > \alpha$ then continue

2. grow merged region from seed region r_i:

   while (1)

   a. find most spectrally similar region connected to r_i:

   $$j \leftarrow \underset{k \in c(r_i)}{\arg\min} \; \| \mathbf{s}(r_i) - \mathbf{s}(r_k) \|^2$$

   b. if $\| \mathbf{s}(r_i) - \mathbf{s}(r_j) \|^2 > \Delta_s^2$ then exit while loop

   c. absorb r_j into r_i

   d. if $area(r_i) > \alpha$ then exit while loop

The process of absorbing r_j into r_i is computationally efficient because region connectivity analysis is performed only once at initialization time. The algorithm then only needs to maintain lists of region ID's:

- $r_i \leftarrow r_i \cup r_j$
- $\mathbf{s}(r_i) \leftarrow \dfrac{area(r_i)\,\mathbf{s}(r_i) + area(r_j)\,\mathbf{s}(r_j)}{area(r_i) + area(r_j)}$
- $area(r_i) \leftarrow area(r_i) + area(r_j)$
- $\forall k \in r_j,\; m_k \leftarrow i$
- $c(r_i) \leftarrow c(r_i) \cup c(r_j)$ with i and j removed
- $\forall k \in c(r_i)$, replace j with i in $c(r_k)$
- $c(r_j) \leftarrow \{\ \}$

Coarseness regulation applies to both monochrome and multi-band images ($\boldsymbol{s}(r_i)$ can be either scalar or vector). It has two parameters: α and Δ_s. Δ_s is not user-specified – its value is based on the limit of human ability to perceive contrast in display brightness (e.g., in 8-bit images, $\Delta_s \approx 4$ is appropriate). In *interactive* coarseness regulation, α is user-specified, whereas in *convergent* coarseness regulation, α is set to the number of pixels in the image. In convergent coarseness regulation (which has *no* user-specified parameters), the output converges to a steady-state segmentation that remains unchanged as α is further increased.

Coarseness regulation can be performed in steps. The *area coarseness regulation step factor* $\Delta_A = 2,3\ldots$ specifies the manner in which the minimum allowable region area increases from step to step. On step k, the minimum allowable region area is $\alpha_k = (\Delta_A)^k \le \alpha$, $k = 1,2\ldots$. *Stepwise coarseness regulation* guarantees segmentations of constant or increasing coarseness for increasing values of α across powers of Δ_A. This intuitively satisfying property is *not* inherent in *one-step coarseness regulation* (for which it is theoretically possible to end up with more regions when α increases). Moreover, for sufficiently small values of Δ_A (say $\Delta_A = 2$ as in Section 4), coarseness regulation is much less likely to suffer from reductions in computational efficiency due to the overwhelming combinatorics associated with pairs of connected regions.

3. Segmentation Quality Assessment

In the absence of ground truth, the best way to build a segmentation reference is for a trained human analyst to use an image annotation tool to draw boundaries of perceived regions with a mouse. Image segmentation quality can then be assessed by measuring the disparity between manually generated and computer-generated region boundaries. It has been empirically established that the most reliable measures of segmentation quality take into account both locations and orientations of boundary pixels.

Consider a manually generated region map with $N_P > 0$ region boundary pixels $\boldsymbol{p} \in P$, and a computer-generated region map with $N_Q > 0$ region boundary pixels $\boldsymbol{q} \in Q$. Then, let $\boldsymbol{p(q)}$ be the pixel $\boldsymbol{p} \in P$ closest to $\boldsymbol{q}$, and let $\boldsymbol{q(p)}$ be the pixel $\boldsymbol{q} \in Q$ closest to $\boldsymbol{p}$. The following measure of disparity $d(P,Q) \in [0,1]$ between P and Q has been found to produce segmentation quality assessments that are highly consistent with human perception ([8]):

$$(1) \qquad d(P,Q) = 1 - \min\left[\frac{1}{N_P}\sum_{\boldsymbol{p}\in P} s(\boldsymbol{p},Q)\,,\ \frac{1}{N_Q}\sum_{\boldsymbol{q}\in Q} s(\boldsymbol{q},P)\right]$$

$$(2) \qquad s(\boldsymbol{p},Q) = \begin{cases} \cos^n\theta\,[\boldsymbol{p},\boldsymbol{q(p)}] & \|\,\boldsymbol{q(p)} - \boldsymbol{p}\,\| \le \Delta \\ 0 & \text{otherwise} \end{cases}$$

$\theta[\boldsymbol{p},\boldsymbol{q}]$ is the *phase disparity* (disparity in boundary direction) between boundaries at pixels $\boldsymbol{p}$ and $\boldsymbol{q}$. It varies from 0 (for angles that point in the same or opposite directions) to $\pi/2$ (for angles that point in orthogonal directions). Quality assessments most consistent with human perception were observed for association distances $\Delta \approx 10$ between pixels $\boldsymbol{p}$ and $\boldsymbol{q}$ combined with phase modulation orders n for which $\cos^n\theta = 0.5$ for θ from roughly $5°$ to $10°$ (i.e., n values from roughly 50 to 200). The experiments in Section 4 use $\Delta = 10$ and $n = 100$.

Note that disparity values range from 0 to 1, but take on non-intuitive absolute values that mostly reflect how the disparity formula parameters were specified. So although absolute disparity values have limited meaning (e.g., a value near 0 does not necessarily indicate a good match, and a value near 1 does not necessarily indicate a poor match), *relative* values of disparity can indeed be used to rank segmentation quality as a function of coarseness.

4. Experiments

Fig.1 shows (a) a rural scene (courtesy of AVIRIS) (b) an aircraft scene (courtesy of DigitalGlobe) and (c) a facility scene (courtesy of IKONOS). The red region boundaries were drawn by a human analyst. The orange region boundaries were generated by computer segmentation and various degrees of coarseness regulation (specifically no coarseness regulation, interactive coarseness regulation with a minimum region area of 100 pixels and convergent coarseness regulation). The images were pre-processed with three, zero and five despeckling iterations respectively ([1]). The pre-processsed images were subjected to K-Means pixel classification ([9]). Over-segmentation was achieved by using $K = 10$. The computer-generated segmentations qualitatively appear to more closely resemble the manual segmentations as coarseness increases. Fig.1(d) shows plots of disparity (d in equation (1)) vs. coarseness (α). The results quantitatively confirm the trend towards smaller disparity between manually generated and computer-generated segmentations as coarseness increases to convergence.

5. Summary and Conclusions

Convergent coarseness regulation has been introduced as a post-processing technique that eliminates the need for human analysts to interactively select values for image segmentation algorithm parameters. It thus enables systems to be developed for which there are requirements to *automatically* segment large amounts of imagery acquired by various imaging sensors at various times. Convergent coarseness regulation results are amazingly insensitive to the initial degree of over-segmentation. However, image pre-processing can still typically have a profound effect on segmentation results.

Selected References

[1] T. Crimmins, “Geometric Filter for Reducing Speckle”, Applied Optics, vol.24, 1985, pp.1438-1443.

[2] C. Boumann and M. Shapiro, “A Multiscale Random Field Model for Bayesian Image Segmentation”, IEEE Trans. Image Proc., vol.3, March 1994, pp.162-177.

[3] T. Pappas, “An Adaptive Clustering Algorithm for Image Segmentation”, IEEE Trans. Signal Proc., vol.40, no.4, April 1992, pp.901-914.

[4] R. Adams and L. Bischof, “Seeded Region Growing”, IEEE Trans. PAMI, vol.16, 1994, pp.641-647.

[5] L. Davis, “A Survey of Edge Detection Techniques”, CGIP, vol.4, 1975, pp.248-270.

[6] S. Osher and J. Sethian, “Fronts Propagation with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations”, Journal of Computational Physics, vol.79, 1988, pp.12-49.

[7] V. Caselles, R. Kimmel and G. Sapiro, “Geodesic Active Contours”, Int. Journal on Computer Vision, February-March 1997, pp.61-79.

[8] D. Paglieroni, “Design Considerations for Image Segmentation Quality Assessment Measures”, Pattern Recognition, Vol.37, No.8, August 2004.

[9] A. Meyer, D. Paglieroni and C. Astaneh, "K-Means Re-Clustering: Algorithmic Options with Quantifiable Performance Comparisons", SPIE Photonics West, Optical Engineering at LLNL, January 2003.

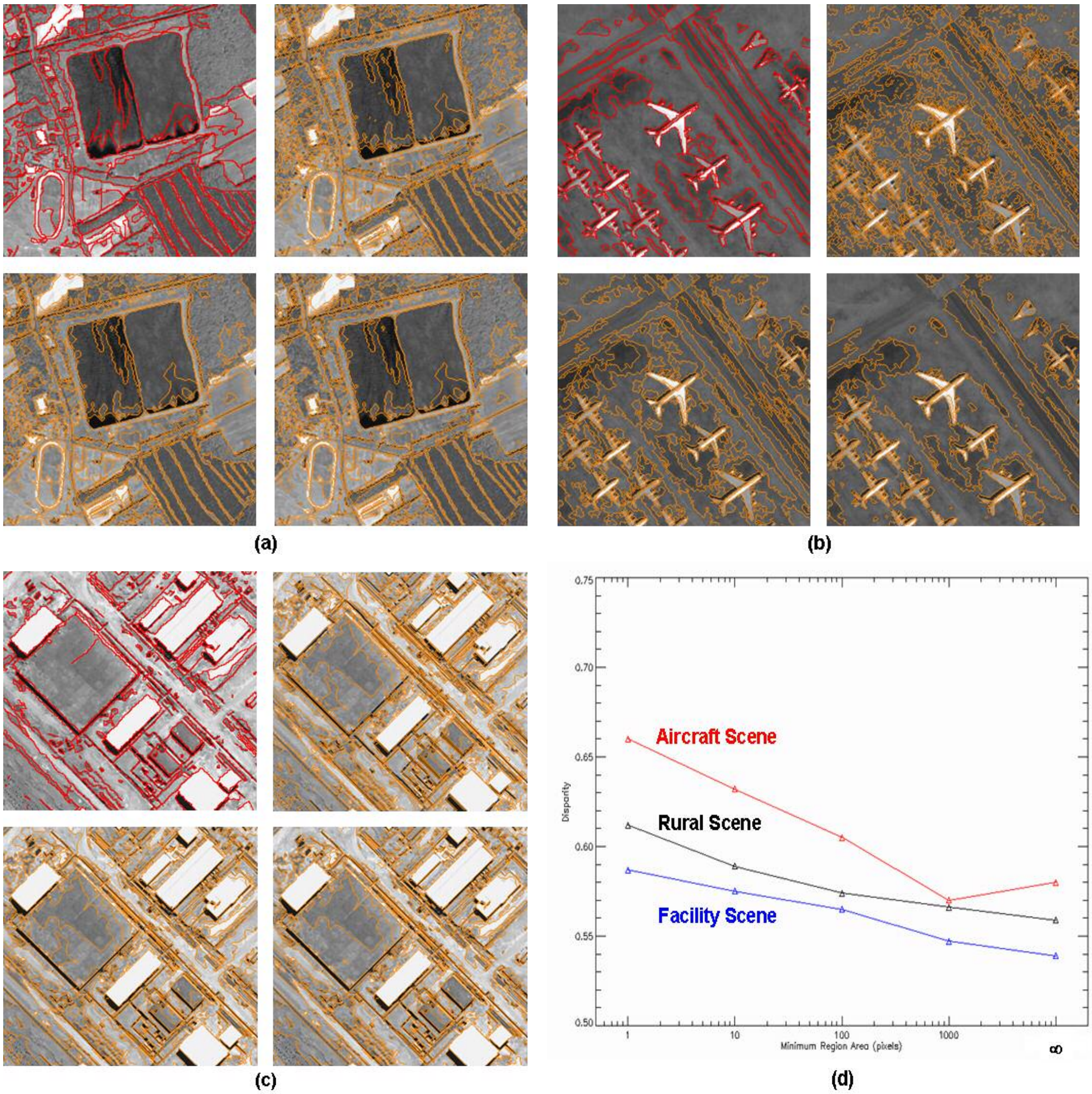

Fig.1 Manual (red) and computer-generated (orange) region boundaries for (a) rural scene (courtesy of AVIRIS) (b) aircraft scene (courtesy of DigitalGlobe) (c) facililty scene (courtesy of IKONOS). Top right – no coarseness regulation. Bottom left – interactive coarseness regulation (minimum region area = 100). Bottom right – convergent coarseness regulation. (d) Plots of disparity between manual and computer-generated segmentations vs. degree of coarseness.